

PERBANDINGAN APLIKASI METODE CROSS ENTROPY DAN PARTICLE SWARM OPTIMIZATION PADA SUPPORT VECTOR MACHINE

Herlina, Dwi Yuli Rakhmawati

Teknik Industri, Universitas 17 Agustus 1945 Surabaya
herlina@untag-sby.ac.id

ABSTRAK

Support Vector Machine (SVM) merupakan salah satu metode klasifikasi yang banyak digunakan oleh para ilmuwan karena memiliki keunggulan dalam menemukan solusi yang bersifat global optimal. Pada penelitian ini, metode *Cross Entropy (CE)* dan *Particle Swarm Optimization (PSO)* akan diterapkan pada SVM untuk permasalahan klasifikasi dua kelas (*binary class classification*). Metode CE dan PSO akan digunakan untuk menemukan solusi optimal dari dual SVM. Dengan diterapkannya metode CE dan PSO pada SVM, akan mempersingkat waktu komputasi jika dibandingkan dengan metode SVM standar dengan tetap mempertahankan tingkat akurasi tetap terjaga dengan baik. Uji coba metode CE-SVM dan PSO-SVM akan diterapkan dengan mengambil dataset dari permasalahan nyata yang diambil dari *UCI repository*, yaitu *Haberman's survival dataset* dan *liver disorders dataset*. Hasil yang didapat dari metode CE-SVM dan PSO-SVM akan dibandingkan dari segi waktu komputasi dan tingkat akurasi. PSO-SVM membutuhkan waktu komputasi yang lebih singkat dibandingkan dengan CE-SVM dan SVM standar. Kedua metode PSO-SVM dan CE-SVM memberikan tingkat akurasi yang baik dibandingkan dengan SVM standar.

Kata kunci: *Cross Entropy, Particle Swarm Optimization, Support Vector Machine, Klasifikasi*

ABSTRACT

Support Vector Machine (SVM) is one of the classification method that many scientists use because it has the advantage in finding solutions that are global optimal. In this research, *Cross Entropy (CE)* and *Particle Swarm Optimization (PSO)* methods will be applied to the SVM for two class classification problems. CE and PSO methods will be used to find the optimal solution of dual SVM. With the adoption of CE and PSO methods on SVM, it will shorten computation time compared to standard SVM methods while maintaining accurate levels of accuracy. Trial of CE-SVM and PSO-SVM methods will be applied by retrieving the datasets from the real problems taken from the *UCI repository*, ie *Haberman's survival dataset* and *liver disorders dataset*. The results obtained from the CE-SVM and PSO-SVM methods will be compared in terms of computation time and accuracy. PSO-SVM requires a shorter computation time compared to CE-SVM and standard SVM. Both PSO-SVM and CE-SVM methods provide a good degree of accuracy compared to standard SVM.

Keywords: *Cross Entropy, Particle Swarm Optimization, Support Vector Machine, Classification*

PENDAHULUAN

Support Vector Machines (SVM) merupakan salah satu metode yang dapat digunakan untuk menyelesaikan permasalahan klasifikasi maupun regresi. Meskipun SVM termasuk metode yang cukup baru namun karena kemampuannya dalam menemukan solusi yang bersifat global optimal, maka SVM banyak digunakan oleh para ilmuwan. SVM sendiri merupakan metode yang sekelas dengan ANN (Artificial Neural Network), namun dalam beberapa penelitian yang telah dilakukan terbukti bahwa SVM memberikan hasil yang lebih baik daripada ANN karena ANN memberikan solusi lokal optimal sedangkan SVM menemukan solusi yang bersifat global optimum (Santosa, 2007).

Seiring dengan semakin banyak penelitian menggunakan metode SVM, menimbulkan munculnya penelitian-penelitian selanjutnya dalam usaha untuk mempersingkat waktu komputasi untuk menemukan solusi optimal.

Ide dasar dari SVM adalah memetakan data input ke dalam ruang berdimensi tinggi dimana nantinya akan dapat ditemukan fungsi pemisah yang linier. Untuk menemukan fungsi pemisah ini, cara kerja SVM adalah memaksimalkan jarak diantara dua titik terdekat dari dua kelas yang berbeda. Hal ini dicapai dengan memformulasikan problem sebagai *quadratic programming*. Untuk menemukan solusi dari permasalahan *quadratic programming* ini diperlukan waktu komputasi yang cukup lama terutama bila jumlah datanya sangat besar (Santosa, 2009). Untuk dapat mempersingkat waktu komputasi dan meningkatkan akurasi dari model maka diperlukan pencarian nilai optimal dari parameter yang digunakan dalam SVM (Herlina, 2012).

Dalam penelitian ini, Cross Entropy (CE) dan Particle Swarm Optimization (PSO) akan diaplikasikan pada *dual* Lagrange SVM untuk mencari nilai optimal atau mendekati optimal untuk salah satu parameter dari SVM yaitu Lagrange multipliers (α) yang selanjutnya akan digunakan dalam SVM. CE adalah salah satu metode optimasi metaheuristik yang awalnya digunakan untuk simulasi kejadian langka namun dalam perkembangannya CE dapat dikembangkan untuk menyelesaikan beberapa kasus machine learning. PSO adalah salah satu metode optimasi yang banyak digunakan karena beberapa kelebihan yang dimilikinya antara lain dalam perhitungannya tidak membutuhkan turunan (*derivative*), dapat keluar dari jebakan lokal optimal, parameter yang digunakan tidak terlalu banyak dan hasil optimal yang ditemukan tidak bergantung pada solusi awal yang dibangkitkan.

Dengan digunakannya metode CE-SVM dan PSO-SVM ini diharapkan akan mempersingkat waktu komputasi dengan tingkat akurasi yang tetap terjaga. Oleh karena itu nantinya akan dibandingkan hasilnya dalam hal waktu komputasi dan akurasi antara metode CE-SVM, PSO-SVM, dan SVM standar.

MATERI DAN METODA

Support Vector Machines

SVM adalah sebuah algoritma yang diusulkan oleh Vapnik pada tahun 1995. SVM tergolong metode klasifikasi baru dan telah banyak dijadikan metode dalam sejumlah penelitian, seperti *pattern recognition*, regresi, dan estimasi. Menurut Santosa (2007), SVM berada dalam satu kelas dengan ANN tetapi dalam banyak implementasi

terbukti bahwa SVM memberikan hasil yang lebih baik daripada ANN dalam hal solusi yang dicapai. ANN menemukan solusi yang berupa lokal optimal sedangkan SVM menemukan solusi yang global optimal.

SVM menggunakan masukan dari data *training* untuk menemukan fungsi pemisah (klasifier/*hyperplane*) terbaik diantara fungsi yang tidak terbatas jumlahnya untuk memisahkan dua macam obyek. *Hyperplane* terbaik didapatkan dengan memaksimalkan margin atau jarak diantara dua set obyek dari kelas yang berbeda. SVM dapat diterapkan pada data yang bersifat linear maupun non-linear. Untuk kasus klasifikasi, dimana datanya tidak linear dapat menggunakan metode Kernel (Han dan Kamber, 2001).

Problem optimisasi dengan menggunakan SVM untuk kasus klasifikasi dengan dua kelas dapat dirumuskan sebagai berikut:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \eta_i \quad (1)$$

Subject to:

$$y_i(w x_i + b) + \eta_i \geq 1, \eta_i \geq 1, i = 1, \dots, m$$

Dimana:

x_i = data *input*

y_i = *output* dari data x_i

w, b = parameter yang akan dicari nilainya

C = parameter yang ditentukan oleh *user* (*penalty error*)

Menggunakan Lagrange *multipliers*, α , *inequality constraints* pada persamaan 1 dapat diformulasikan sebagai dual Lagrange sebagai berikut:

$$L(\alpha) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^m \alpha_i \quad (2)$$

Subject to:

$$\sum_{i=1}^l \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C, i = 1, \dots, m$$

Dimana m adalah jumlah data yang digunakan untuk *training*.

Decision function yang akan dihasilkan memenuhi rumusan:

$$f(x) = \text{sign}(\sum_{i \in SV} y_i \alpha_i K(x, x_i) + b) \quad (3)$$

Untuk mengatasi permasalahan yang bersifat tidak linier, dapat digunakan metode kernel. Dengan metode kernel suatu data x di *input space* dimapping ke *feature space* F dengan dimensi yang lebih tinggi. Suatu kernel map mengubah problem yang tidak linier menjadi linier dalam space baru. Fungsi kernel yang biasanya dipakai dalam literatur SVM (Haykin, 1999 dalam Santosa, 2007):

1. Linear: $x^T x$
2. Polynomial: $(x^T x_i + 1)^p$
3. Radial basis function (RBF): $\exp(-\frac{1}{2\sigma^2} \|x - x_i\|^2)$
4. Tangent hyperbolic (sigmoid): $\tanh((\beta x^T x_i + \beta_i))$, dimana $\beta, \beta_i \in R$

Pemilihan jenis fungsi kernel yang akan digunakan untuk substitusi *dot product* di *feature space* akan sangat bergantung pada data.

Cross Entropy

Metode CE merupakan metode yang cukup baru yang awalnya diterapkan untuk simulasi kejadian langka (*rare event*) kemudian berkembang untuk beberapa kasus antara lain optimasi kombinatorial, optimasi kontinyu, dan *machine learning*. Metode CE termasuk dalam teknik Monte Carlo yang bisa digunakan untuk menyelesaikan kasus optimasi. CE dapat digunakan untuk menyelesaikan permasalahan optimasi kombinatorial yang kompleks dengan cara meminimalkan cross entropy yang dilakukan dengan cara menerjemahkan masalah optimasi deterministik menjadi stokastik, kemudian menggunakan teknik simulasi kejadian langka (Santosa, 2015).

Dalam metode CE ada aturan penting untuk mengupdate parameter. Ide utama dari metode CE dapat dinyatakan sebagai berikut: misalnya terdapat suatu masalah untuk meminimalkan suatu fungsi $f(x)$ pada setiap x yang berasal dari χ dimana nilai minimum yang didapat adalah Y^* ,

$$Y^* = \min_{x \in \chi} f(x) \quad (4)$$

Kita perlu melakukan beberapa langkah untuk menemukan x sehingga $f(x)$ minimum. Pertama kita bangkitkan bilangan random x melalui suatu *probability density function* (pdf) tertentu. Misalnya, bangkitkan nilai x yang berdistribusi normal sejumlah N sampel. Untuk distribusi normal, diperlukan parameter nilai μ dan σ untuk membangkitkan χ . Lalu tentukan parameter ρ yang tidak terlalu kecil, misalnya $\rho = 0,1$. Parameter ini menentukan berapa persen dari seluruh sampel yang akan kita gunakan untuk mengupdate parameter v berikutnya. Nilai ρ akan menentukan berapa banyak dari N sampel, porsi yang akan diambil sebagai sampel elite. Selain itu diperlukan konstanta α yang digunakan untuk membobot parameter pada iterasi sekarang dan iterasi sebelumnya. Metode CE melibatkan prosedur iterasi, dimana tiap iterasi dapat dipecah menjadi dua fase:

- a. Membangkitkan sampel random (x) dengan menggunakan mekanisme atau distribusi tertentu.
- b. Memperbaharui parameter (v) dari mekanisme random berdasarkan data sampel elite untuk menghasilkan sampel yang lebih baik pada iterasi berikutnya.

Particle Swarm Optimization

PSO diusulkan oleh Kennedy dan Eberhart pada tahun 1995. Metode PSO didasarkan pada perilaku sekawanan burung atau ikan. Algoritma PSO meniru perilaku sosial organisme ini yang terdiri dari tindakan individu dan pengaruh dari individu-individu lain dalam suatu kelompok (Santosa dan Willy, 2011).

Pada algoritma PSO ini, pencarian solusi dilakukan oleh suatu populasi yang terdiri dari beberapa partikel. Populasi dibangkitkan secara random dengan batasan nilai terkecil dan terbesar. Setiap partikel merepresentasikan posisi atau solusi dari permasalahan yang dihadapi. Setiap partikel melakukan pencarian solusi yang optimal dengan melintasi ruang pencarian (*search space*). Hal ini dilakukan dengan cara setiap partikel melakukan penyesuaian terhadap posisi partikel terbaik dari partikel tersebut (*local best*) dan penyesuaian terhadap posisi partikel terbaik dari seluruh kawanan (*global best*) selama melintasi ruang pencarian. Setelah itu dilakukan proses pencarian untuk mencari posisi terbaik setiap partikel dalam sejumlah iterasi tertentu sampai didapatkan posisi yang relatif *steady* atau mencapai batas iterasi yang telah ditetapkan. Pada setiap iterasi, setiap solusi yang direpresentasikan dengan posisi partikel, dievaluasi performansinya dengan memasukkan solusi tersebut ke dalam *fitness function* (Santosa dan Willy, 2011).

Formulasi matematika yang menggambarkan posisi dan kecepatan partikel pada suatu dimensi ruang tertentu dapat dituliskan sebagai berikut:

$$X_i(t) = x_{i1}(t), x_{i2}(t), \dots, x_{iN}(t) \quad (5)$$

$$V_i(t) = v_{i1}(t), v_{i2}(t), \dots, v_{iN}(t) \quad (6)$$

Dimana

X = posisi partikel

V = kecepatan partikel

i = indeks partikel

t = iterasi ke-t

N = ukuran dimensi ruang

Berikut ini merupakan model matematika yang menggambarkan mekanisme updating status partikel Kennedy dan Eberhart (1995):

$$V_i(t) = V_i(t-1) + c_1 r_1 (X_i^L - X_i(t-1)) + c_2 r_2 (X^G - X_i(t-1)) \quad (7)$$

$$X_i(t) = V_i(t) + X_i(t-1) \quad (8)$$

Dimana

$X_i^L = x_{i1}^L, x_{i2}^L, \dots, x_{iN}^L$ merepresentasikan *local best* dari partikel ke-i.

$X^G = x_1^G, x_2^G, \dots, x_N^G$ merepresentasikan *global best* dari seluruh kawanan.

C_1 dan c_2 adalah suatu konstanta yang bernilai positif dan biasanya disebut sebagai *learning factor*.

r_1 dan r_2 adalah suatu bilangan random yang bernilai antara 0 sampai 1.

Algoritma PSO meliputi langkah berikut (Santosa dan Willy, 2011):

Pertama, bangkitkan posisi awal sejumlah partikel sekaligus kecepatan awalnya secara random.

Selanjutnya, evaluasi *fitness* dari masing-masing partikel berdasarkan posisinya.

Kemudian, Tentukan partikel dengan *fitness* terbaik, dan tetapkan sebagai *Gbest*. Untuk setiap partikel, *Pbest* awal akan sama dengan posisi awal.

Ulangi langkah berikut sampai *stopping criteria* dipenuhi.

Pertama, Menggunakan *Pbest* dan *Gbest* yang ada, perbarui kecepatan setiap partikel menggunakan persamaan (7). Lalu dengan kecepatan baru yang didapat, perbarui posisi setiap partikel menggunakan persamaan (8).

Kedua, evaluasi *fitness* dari setiap partikel.

Ketiga, tentukan partikel dengan *fitness* terbaik, dan tetapkan sebagai *Gbest*. Untuk setiap partikel, tentukan *Pbest* dengan membandingkan posisi sekarang dengan *Pbest* dari iterasi sebelumnya.

Cek *stopping criteria*. Jika dipenuhi, berhenti. Jika tidak, kembali ke langkah pertama.

Metode Penelitian

Penelitian ini dilaksanakan dengan tahapan sebagai berikut:

Pertama, membuat algoritma metode CE-SVM.

Kedua, membuat algoritma metode PSO-SVM.

Ketiga, melakukan uji validasi terhadap metode CE-SVM dan PSO-SVM.

Keempat, menetapkan dataset yang akan digunakan sebagai uji coba metode CE-SVM dan PSO-SVM.

Dataset yang dipilih adalah dataset untuk kasus klasifikasi dua kelas, yaitu *Haberman's survival dataset* dan *liver disorders dataset*.

Melakukan uji coba terhadap metode CE-SVM dan PSO-SVM dengan menggunakan dataset dari UCI repository.

Uji coba validasi metode CE-SVM dan PSO-SVM pertama kali diterapkan untuk kasus klasifikasi dua kelas.

Membandingkan hasil dari CE-SVM, PSO-SVM, dan SVM standar untuk dataset awal maupun untuk dataset yang telah dilakukan seleksi fitur dari segi waktu komputasi dan tingkat akurasi.

HASIL DAN PEMBAHASAN

Setelah dibuat kode (*coding*) untuk metode SVM untuk dua kelas, kemudian parameter Lagrange Multiplier dari SVM untuk dua kelas tersebut dioptimalkan menggunakan PSO dan CE.

Data yang akan digunakan untuk uji coba model menggunakan dataset yang diambil dari UCI repository. Data yang dipilih adalah dataset jenis klasifikasi dua kelas (*binary*), yaitu *Haberman's survival datasets* dan *liver disorders datasets*. *Haberman's survival datasets* memiliki 3 atribut dengan 306 data, sedangkan untuk *Liver disorders datasets* memiliki 6 atribut dan 345 data.

Dataset yang dipilih akan dibagi menjadi *data training* dan *data testing*. Pembagian jumlah data *training* dan data *testing* untuk masing-masing dataset mengikuti rasio 2:1.

Uji coba akan dilakukan sebanyak lima kali dengan kombinasi data *training* dan data *testing* yang berbeda untuk masing-masing dataset. Tingkat akurasi dan lama waktu komputasi dari kelima kombinasi tersebut akan dirata-rata untuk menemukan satu hasil.

Data *training* dan data *testing* yang telah ditentukan akan diuji dengan menggunakan *software* MATLAB. Kernel yang akan digunakan dalam penelitian ini adalah RBF dengan $\sigma = 5$ untuk SVM standar. Sedangkan untuk metode PSO-SVM, kernel yang akan digunakan adalah RBF dengan $\sigma = 5$, nilai Lagrange Multipliers yang akan dibangkitkan sejumlah 10, dan maksimum iterasi 50. Untuk metode CE-SVM, kernel yang akan digunakan adalah RBF dengan $\sigma = 7$, jumlah random sampel yang akan dibangkitkan sejumlah 20, dengan $\rho = 0,2$ dan parameter smoothing 1. Uji coba dilakukan menggunakan prosesor komputer Intel Pentium CPU B960 2.2 GHz, 2 GB RAM.

Hasil dari uji coba permasalahan klasifikasi dua kelas dengan menggunakan metode SVM, PSO-SVM, dan CE-SVM ditampilkan pada Tabel 1.

Tabel 1. Hasil Perbandingan Klasifikasi Dua Kelas

Data	SVM Standar		PSO-SVM		CE-SVM	
	Waktu Komputasi (detik)	Misklasifikasi	Waktu Komputasi (detik)	Misklasifikasi	Waktu Komputasi (detik)	Misklasifikasi
<i>Haberman's survival</i>	2,27	31,7%	0,68	25,7%	0,79	26,5%
<i>Liver disorders</i>	2,30	34,8%	0,80	31,6%	0,98	30,4%

Dari Tabel 1 menunjukkan bahwa metode optimasi PSO dan CE yang diaplikasikan pada SVM dapat mempersingkat waktu komputasi dan memberikan tingkat akurasi yang cukup baik dilihat dari semakin kecilnya persentase misklasifikasi yang terjadi. Misklasifikasi adalah tingkat kesalahan dalam mengklasifikasikan data ke dalam kelasnya. Metode PSO-SVM dalam penelitian ini memberikan waktu komputasi yang lebih singkat dibandingkan CE-SVM. Waktu komputasi ini juga dipengaruhi oleh pemilihan parameter-parameter yang digunakan baik pada metode PSO-SVM maupun CE-SVM.

KESIMPULAN

Dari hasil penelitian yang telah dilakukan dapat disimpulkan antara lain:

1. Metode PSO dan CE dapat diterapkan untuk menemukan solusi optimal dari permasalahan dual Lagrange SVM..
2. Metode PSO-SVM dan CE-SVM dapat menyelesaikan kasus klasifikasi dua kelas dengan waktu komputasi yang lebih singkat dibandingkan dengan SVM standar.
3. Tingkat akurasi yang diperoleh menggunakan metode PSO-SVM dan CE-SVM cukup baik dibandingkan dengan metode SVM standar. Tingkat akurasi yang diperoleh menggunakan metode PSO-SVM dan CE-SVM cukup baik dibandingkan dengan metode SVM standar.

DAFTAR PUSTAKA

- Han, J. dan Kamber, M., 2001, *Data Mining: Concepts and Techniques*, 2nd edition, Elsevier Inc., San Fransisco.
- Herlina, 2012, *Pemodelan Prediksi Finansial Distress Menggunakan Metode Particle Swarm Optimization-Support Vector Machine*, Thesis Program Magister Bidang Optimasi Sistem Industri Jurusan Teknik Industri Institut Teknologi Sepuluh Nopember.
- Santosa, B., 2007, *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Graha Ilmu, Yogyakarta.
- Santosa, B., 2009, Application of the Cross-Entropy Method to Dual Lagrange Support Vector Machine, *Lectures Notes in Artificial*, Springer.
- Santosa, B. dan Willy, P. 2011, *Metoda Metaheuristik Konsep dan Implementasi*, Guna Widya, Surabaya.
- Santosa, B., 2015, Multiclass Classification with Cross-Entropy-Support Vector Machines, *The Third Information Systems International Conference*, 72(2): 345-352.